

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 23:36:15 ; Search time 5307.2 Seconds
(without alignments)
2746.592 Million cell updates/sec

Title: US-09-719-748-1_COPY_62_1141

Perfect score: 1080
Sequence: 1 atgagcgcattcaacagcagca.....ggagggagagcagcacctcc 1080

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	781	72.3	1729	11	AK004350	AK004350 Mus muscu
2	604.2	55.9	882	9	AI322362	me98c10.y
3	532.6	49.3	830	10	BG68701	602787522
4	512.8	47.5	976	11	BC012210	Mus muscu
5	506.6	46.9	1199	11	AK009701	Mus muscu
6	506.4	46.9	940	10	BE883503	601511064
7	499.2	46.2	922	10	BT554559	603235693
8	495.6	45.9	649	10	BT046467	MR3-FN020
9	486.8	45.1	555	10	BF087508	QV2-HT054
10	459.2	42.5	869	10	W82116	me98c10.y
11	453.4	42.0	585	10	BG384579	303638 MA
12	434.6	40.2	565	10	BF075625	224820 MA
13	431.6	40.0	999	10	BF168866	601775325
14	429.8	39.8	1127	10	BF733365	601567911
15	428.2	39.6	671	10	BF727181	by17h06.y
16	425.6	39.4	556	10	BF075639	224840 MA
17	425.6	39.4	982	10	BF976537	602244391

18	415.4	38.5	708	9	AM476323	uq3f11.y
19	411.2	38.1	696	10	BF012103	ux53g01.y
20	409	37.9	1058	10	BG421064	602451186
21	408.4	37.8	460	9	AA858002	0163g08.s
22	408.2	37.8	663	10	BF019568	ux12f12.y
23	402.8	37.3	681	10	BF149425	602848986
24	402	37.2	481	9	AM603538	RCO-CN002
25	402	37.2	887	10	BG968191	602835674
26	397.8	36.8	871	10	BF690902	603312045
27	397.4	36.8	855	10	BF182782	603074924
28	394.2	36.5	703	10	BG65492	602783624
29	393	36.4	756	10	BF162798	601769262
30	388.8	36.0	796	10	BG421646	602449739
31	386.8	35.8	797	10	BF163166	603286787
32	380.6	35.2	710	10	BG277312	ux42b06.y
33	377	34.9	1039	10	BG419640	602451855
34	375.6	34.8	785	10	BF152999	603299971
35	374.2	34.6	707	10	BF1904738	603168651
36	371.4	34.4	654	10	BF012312	BF012312
37	370.6	34.3	441	9	AA718515	VR84F03.F
38	369	34.2	798	10	BG662263	602795580
39	364.2	33.7	693	9	AL631466	AL631466
40	355.6	32.9	743	10	BG968192	6028356
41	355.4	32.9	958	10	BF157256	603030761
42	354.6	32.8	623	10	BE848999	uw01907.y
43	354.4	32.8	622	10	BF1414743	602991133
44	353.2	32.7	433	10	W34247	mb80c12.t1
45	352.2	32.6	548	10	BF558103	BF558103 UT-R-C0-h

ALIGNMENTS

RESULT	1	1729 bp	MRNA	linear	HTC 19-JAN-2002
AK004350	AK004350	1729 bp	MRNA	linear	HTC 19-JAN-2002
LOCUS	AK004350	1729 bp	MRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus 18 days embryo whole body cDNA, RIKEN full-length enriched library, clone:1110062102:death-associated kinase 2, full insert sequence.				
ACCESSION	AK004350	GI:12835501			
VERSION	AK004350.1	GI:12835501			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (Strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:1110062102.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (sites)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
AUTHORS	2 (sites)				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normaliation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	3 (sites)				
AUTHORS	Shibata, K., Itch, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashito, H., Itch, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system-384 format				
JOURNAL	sequencing pipeline with 384 multicapillary sequencer				
	Genome Res. 10 (11), 1757-1771 (2000)				

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:247370
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 442.
Location/Qualifiers
1. .882

Query Match	55.9%	Pred. 604.2;	DB 9;	Length 882;
Best Local Similarity	87.7%	Pred. No. 4.3e-129;		
Matches 690;	Conservative	0;	Mismatches 94;	Indels 3;
				Gaps 3

QY	121	CCGGAGGAGAACTGTTTGGATTCTCTGGCCACAGAGAGTCGTTAATGATGAGGAGAAAGCA	180
QY	350	ccagcttcataagcgcagatctctgtagaggggtgaacactctcaacaaagaanaattgtc	409
Db	181	CCAGCTTCATTAAAGCAGATCTCTGGATGGGGTGTAATTACCTTACACAAAGAAATTTGCTC	240
QY	410	actttagtcacagccagaacacatattgtgttagacaagaattcccatccaaca	469
Db	241	ACTTTGATCTCAAGCCAGAAAACATATATGTGTGTAGACAGAAATATCCCAATTCACACA	300
QY	470	tcaactgtatgtacttgtgtctgtgtccacagaataagatgtagtgaattgaattaaaga	529
Db	301	TCAAGCTGATGTAGCTTTGGCTTGCTGCACAGAAATAGAAATGATGAGTGAATTAAATAACA	360
QY	530	tttttggagcccggaatttgtgtgtccacagaataattgtgaactcaagagccctgggtctgg	589
Db	361	TTTTTTGGGACACCTGGAATTTGTTGGCTCCAGAAATCGTGAACATATAGCCACATGMCATGG	420
QY	590	agagctgacatgctgagatagcgtcatcaaccatacctctctaagttagagatccctt	649
Db	421	AGGCCGACATGTGGAGCATTTGGAGTATCATCACTATATTTCTTTAGTGGAGCGTCCC-C	479
QY	650	tccttgtagagacacgaagacagagaacacatggtcaataatcacaatcagtgagttacgattg	709
Db	480	TCCTGGGAGACACAGAAACAGAAACCTGGCTAATATCATCTGCTGTGAGTATTCAGCATMTG	539
QY	710	atgaggaattcttcacacatacgagcgagctgtgccaagacttattcggaaactcttgg	769
Db	540	ATGAGGAAATTTCTTCAGCCACACAGAGAGAGCTGGCCCAAGACATTCATTTGMAAGCTCTTG	599
QY	770	ttaaagagaccgggaagcgctcacatctcaagatcaagagctctagaacccctgtgacagc	829
Db	600	TGAAGAGACCCCGGAANCGGCTTACCATCATCAAGAGGCTCTCAGACATTCCTGTGATCAGC	659
QY	830	cgtgtgacacacacagacagcatgtgtgacagggatctgtgtcaatctggaagactca	889
Db	660	CCGTGGACACCCAGCAAGCATTTGTATGCAGAGAGTCTGTGTCAACCTGAGGAATTTTA	719
QY	890	ggaagcagatgtctcgcagcggtgtgaaagcttcccttcacacatcgtgtccctgtgcaac	949
Db	720	GAAAGCAG-ATGTTCCGAAGCCGGGAGAGCTTGNCCTTAGCATCGGCTCCTTG-CAAC	777
QY	950	accctcac 956	
Db	778	ACCTAAC 784	
RESULT	3		
LOCUS	BG868701	830 bp	mRNA linear EST 29-MAY-2000
DEFINITION	602787522F1 NCI_GCAP_SG2 Mus musculus cDNA clone IMAGE:4913788		
ACCESSION	BG868701		
VERSION	BG868701.1	GI:14219241	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus		
JOURNAL	1 (bases 1 to 830)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Jeffrey E. Green, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://lmage.lnl.gov		
	Plate: LLAM10819 row: k column: 05		

High quality sequence stop: 826.
Location/Qualifiers
1. .830
/organism="Mus musculus"
/strain="FVB/N"
/db.xref="taxon:10090"
/clone="IMAGE:4913788"
/lab_host="NCL CGAP_SG2"
/note="Organ: salivary gland; Vector: pcwv-sport6; Site: 1;
Noti: Site 2: Salt: Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCL CGAP Library."
BASE COUNT 217 a 205 c 244 g 163 t 1 others
ORIGIN

Query Match 49.3%; Score 532.6; DB 10; Length 830;
Best Local Similarity 90.6%; Pred. No. 1.4e-112;
Matches 623; Conservative 0; Mismatches 60; Indels 5; Gaps 5;

Qy 1 atggagccattcaagcagaaggtgagagactttatagacatcgagagagctgggg 60
Db 143 ATGAGAGAGCTTCAAAACAGACAGAGGTGAGACTTTATGACATGAGAGAGAGTGGGC 202

Qy 61 agtggccagtttgcacatcgtaagaagtgcgggagagaagcaagggtctgagatgca 120
Db 203 AGTGCGCCAGTTTGCATCGTGAAGAGAGCGCGGAGAGACACAGGCGCTGAGATGCA 262

Qy 121 gccaaattcatcaagaagcagagccggcgagccggcggtgtgtagccggagagag 180
Db 263 GCCAAGTTTCATTAAG 322

Qy 181 atcgagcgagaggtgagatccttcgagaggtgctgcacacaaatgcatcagctgcac 240
Db 323 ATGAGCGGAGGTGAGAGATCTCGCGGAGAGTGTGACACCCACATCATCAGCTGCAC 382

Qy 241 gacgtctatgagaacccgagcagctggtgacatccttgagctagtgctcggagagag 300
Db 383 GACGTCTATGAGAACCCAGACGAGCTGTGCTCATCTTGAAGTAACTCCGAGAGAGAA 442

Qy 301 cctctcgaattcctcgccagagaagagctgagtgagagagggagccagagctcat 360
Db 443 CTGTTTGAATTCCTGCGCCACAGAGAGTCTTAAGTGAAGAGAGAGAGAGAGAGAG 502

Qy 361 aagcagatcctgagatggggtagaaccttcacacaaagaatgctcacttgatctc 420
Db 503 AAGCAGATCCTGAGTGGGTGAATTAATTCCTCACAAAGAAATGGCTCAGCTTGATCTC 562

Qy 421 aagcagaatacatatggttta-gacaagaatatcccatccacatcaagctgat 479
Db 563 AAGCAGAGAAACATCATGTTGTAGACAGAAATATCCCATTCACATCAAGCTGAT 622

Qy 480 tgaatttgctgcacagaaatagaagatgga-gttgaatttaagatatc-cttggg 537
Db 623 TGATTTGGCGCTGCTCAGCAAAATAGAAGATGAGAGCTTGAATTAATAAATGTTGGG 682

Qy 538 aagcaggaattgtgctcagaatagtgaactaagagccctggg-tctggaagctga 596
Db 683 ACACCTTAATTTGTTGCTCCGAAATCGTGAATATGATGAGCCAGCACTGGAGTGGCA 742

Qy 597 catgtgagatagaggtgcatcacctacatcctcttaagttagagatccctctcctgg 656
Db 743 CATGTGAGCGCTTGGAGTCATCACCTTAATCTTTAGTGGAGCGTCCGCTTCCCTGGG 801

Qy 657 agacagagagagaaacactgcaaat 684
Db 802 AGACACAAACAGAAACCTTGCAAAATT 829

RESULT 4
BC012210 976 bp mRNA linear HTC 08-AUG-2001
LOCUS BC012210 976 bp mRNA linear HTC 08-AUG-2001

DEFINITION Mus musculus, Similar to Death-associated like kinase, clone
IMAGE:3993055, mRNA.
ACCESSION BC012210
VERSION BC012210.1 GI:15126556
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 976)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lochar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.R., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Munry, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAC Plate: 23 Row: n Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomscan gene
prediction, similarity but not identity to protein
This clone has the following problem: Incomplete processing.
Location/Qualifiers
1. .976
/organism="Mus musculus"
/db.xref="taxon:10090"
/clone="IMAGE:3993055"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_id="NCL CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pcwv-sport6"
BASE COUNT 240 a 271 c 294 g 171 t
ORIGIN

Query Match 47.5%; Score 512.8; DB 11; Length 976;
Best Local Similarity 76.2%; Pred. No. 5.7e-108;
Matches 631; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1 atggagccattcaagcagaaggtgagagactttatagacatcgagagagctgggg 60
Db 95 ATGTCACATTCACGACAGAGAGATGTTGAGACCAATTATGAGATGGAGAGAGCTTGGC 154

Qy 61 agtggccagtttgcacatcgtaagaagtgcgggagagaagcaagggtctgagatgca 120
Db 155 AGTGCGCCAGTTTGCATCGTGGCGCAAGTGCAGCAAGAGGCGAGGAGTGAAGTGGCA 214

Qy 121 gccaaattcatcaagaagcagagccggcgagccggcggtgtgtagccggagagag 180
Db 215 GCCAAGTTTCATTAAG 274

Qy 181 atcgagcgagaggtgagatccttcgagaggtgctgcacacaaatgcatcagctgcac 240
Db 275 ATGGAACGCGAGGTGAGATCTCGCGGAGAGATCCGACCCACACATCATTAACACTGAT 334

Qy 241 gacgtctatgagaacccgagcagctggtgacatccttgagctagtgctcggagagag 300

Yamanoto, R., Matsumoto, H., Sakaiguchi, S., Ikegami, T., Kasaiyagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Tonedake, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN Integrated sequence analysts (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
11076861
4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1199)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schirml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toy, T., Yamamura, T., Yamahata, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physicall and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL: <http://genome-gsc.riken.go.jp/>, Fax:81-45-503-9216)

Please visit our web site (<http://genome-gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGTAAATTAATATGCCCCCCCCCCC 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCCAGTAAATTAATATGCCCCCCCCCCC 3']. cDNA was clea with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

Location/Qualifiers

1. 1199
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGI:1901785"
/db_xref="taxon:10090"
/clone="2110039H24"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
165. >1199
/note="DEATH ASSOCIATED PROTEIN KINASE data source:SPR, source key:Q9J9J7, evidence:ISS putative"
/codon_start=-1
/protein_id="BA66448.1"
/db_xref="GI:1284458"
/translation="MTVFREQNDVDYDTGEEGSGGFVAVKKCRKSTGLQVAKFTIKRRTSNRGCVTRDELRVSLIKELRHNVNITLHEVYENKDDVILILEVAGGELEF

DFLAEKESLITEEATEPFLKQILSGVYLHSIQIAHFIDUKENIMLIDRNVKPRIKII
 DFGIAHKIDGNEKFNIGTPEVAPAEIVNEPEKIVLADMSIVITVYIILSGASPL
 GDFKOFELAVNSAVNDEEEFEFPSTLAKDFIRLLVDPKRMJIOSLOHPWIK
 PKDFOALSRRKASAVNNEKFRFARKKMVRICGHIIEFPCVLSEPSNSNGNE
 IYSPNNKKVHC"
 polyA_signal
 1182..1187
 /note="putative"
 1199
 polyA_site
 /note="putative"
 330 a 271 c 342 g 256 t
 BASE COUNT
 ORIGIN
 Query Match 46.9%; Score 506.6; DB 11; Length 1199;
 Best Local Similarity 71.9%; Pred. No. 1.7e-106;
 Matches 662; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

10 ttcaagcagcgaaggtgtaggactttatgacatcgagagagctgaggagtgccag 69
 174 TTCAGGCGAGAAACGTGACGACTACTACGACACCGCGAGAGAACTGGCGAGTGGACAG 233
 70 ttgcacatcgtagaagatgcccggagagagacaggggcttgatgacgccaagtcc 129
 224 TTGCGAGTTGAGAAATGTCGTAGAAAAGTACCGGCTTTCATATCGCGCCAGATTG 293
 130 atcaagagcggcagagccggcgagcgcggtgtgagccggagagagatcgagcg 189
 294 ATCAAGAAAGAGAGAGACCAAGTCACGCGCGGCGGTGACGCGGAGAGACATCGAGCGG 353
 190 gagtgagacatctctgcccaggtgtgtacacacaaatgtacatgacgtgacagactat 249
 354 GAGGTGACATCTCTTAAGAGATCCGACACCAAAATGTCATACCTCGATGAGGTCTT 413
 250 gagaacccagcagcgtgtgacacatcttgagctagtgctgtagagagagctctcgat 309
 414 GAGAACAGACAGATGATCTATCTGATCCTGGAGCTGTGTCAGAGAGTGGAGCTTTGAC 473
 310 ttctgtgcccgaagagtagctacgtgtagagagagagccacagcttcaataagcagatc 369
 474 TTCTGCGCTGAGAGAGAAATCTCTGACTAAGAGAGAGACACGAAATCTCTTAAGCAGATT 533
 370 ctgagtgaggtagaactcttcacacaaagaatgtctcacttgatctcagcagagaa 429
 534 CTCAGCGGCGTTTACTACTGACACTGACATGCGCTCACTTGTGACCTGAAGCCGGAA 593
 430 aacatagtggttagaagaagatatcccatccacacataagctgattgacttggt 489
 594 AACATATGCTTCTGATAGAAATGTGCCAANACCTGGATCAAGATCATAGACTTTGGC 653
 430 ctgagtcacgaataagagatgagtggaattgaataatttttggagcgcggaatt 549
 654 TTGGCGCCATAAATTGACTTTTGAATGAAATCAAAACATATTTGGGACACACAGATT 713
 550 gttgtctcagaagaattgtgaactcagagcccggtgtctgagagctgagactgtggagcata 609
 714 GTGGCTCGGAGATGTCAACTATAGCCCTGGGTCTTGAAGGCGAGATATGTGGAGCATC 773
 610 ggcgtcatcacatcatctcttaagttagagcatcccttctctgtagagacagagcag 669
 774 GGGGTAAATACCTATATCTCTCTAAGTGGGCGCTCCCTTTCTTGGAGACACAGCAA 833
 670 gaaacacitggcaaatatcacatcagtgagttacgaacttgatgaggaatttctcagccat 729
 834 GAATCATTTAGCAATGTGTCGCTGTCAACTACGACTTTGAAGAGAAATTTCTTCGGAAC 893
 730 acgagcagagcttggccaagaactttatcgggaagcttctgtttaaagagagacccggagaag 789
 894 ACCATACCTCTTGGCAAGATTTCATCAGAAAGATGCTGGTCAAGATCCAAAGAGAG 953
 790 ctcaaatcccaagagctctcagacaccccttgatcagcgcgtgtgacaacagcaagcc 849
 954 ATGACAAATCCAGACAGTTTGCAGACACCCCTGTGATCAAGCCTTAAGAGACACCAACAGCA 1013

850 atgtgcagcggagatctgtgtgtaactctgagaaactcaggaacagatctccgag 909
 1014 CTTACTCGAAAGAGCTCAGCATTAACATGAGAAATTCAGAGGTTTCACCTCGAATA 1073
 910 cgtgtgaagcttctcctcagc 930
 1074 AAATGGAAGTAAAGACTCTGC 1094
 Db 1074 AAATGGAAGTAAAGACTCTGC 1094
 RESULT 6
 BE883503
 LOCUS
 DEFINITION
 BE883503
 mRNA sequence.
 BE883503
 VERSION
 KEYWORDS
 EST.
 GI:10332279
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 940)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 TISSUE
 Tissue Procurement: ATCC
 CDNA LIBRARY
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA SEQUENCING
 DNA Sequencing by: Incyte Genomics, Inc.
 CLONE DISTRIBUTION
 MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM9731 row: 0 column: 15
 High quality sequence stop: 643.
 Location/Qualifiers
 1..940
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3912806"
 /clone_11b="NIH-MGC_71"
 /tissue_type="telomerosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pCMV-Sport6; Site: 1; Nci;
 /site_2: Salt; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb.
 BASE COUNT 224 a 263 c 280 g 173 t
 ORIGIN
 Query Match 46.9%; Score 506.4; DB 10; Length 940;
 Best Local Similarity 98.5%; Pred. No. 1.7e-106;
 Matches 532; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

535 gggagccggaaattgtgtctccagaatgtgaaatacagagcccttggtcggagct 594
 1 GGGAGCCCGGAATTTGTCCTCCAGAAATGTGAATGACAGCCCTGGGTGAGGCT 60
 595 gacatgtgagatagagcgtcatcaactacatcctttaaagtggagacatcccttcctg 654
 61 GACATGTGAGATAGGCTCTCATCACTACATCTCTTAAAGTGGAGATCCCTTCTCG 120
 655 gagaacacgaacagaaacacttggcaaatatcacatcagtgagttagacttgatgag 714
 121 GGAGACACGAACAGGAACAACACTGGCAATATACAGCAGTGTGATGAG 180
 715 gaattcttcagcacaagcagcagctgtgccaagagacttattcogaagcttctgtttaa 774
 181 GAATCTTTCAGCCAGACGAGCTGCGCAAGAGATTATTCGGAAGCTTCTGTGTTAA 240
 775 gagaccggaaacggctcacaatccaaagaggtctcagacaccccggtctcagcgggtg 834

OY	128	tcatcaaaagacggaagaagccggaggaccggcgsgfgrttagccggaggaagatctcagc	187
Db	123	TCATCAAAAGACGGGCGCTGCATCCACC6CGGGGTGTGACCGGGAGAATCTTGAAAC	182
OY	188	ggagagtgaagatccttgcgaggtgtgcctcacccaattgatcaagctgcagcgtct	247
Db	183	GGAAGTGAAGCATCTCGCGCAGATCCGCCAACCCACATCATATACTGATGACTGT	242
OY	248	atggaacgcagacogaatgtgtgcacalcctttagactagtltctggaagagaagctctcg	307
Db	243	TGCGAACAACAACAGANTGTGGTGTGATCTCTGAGAGCTGTGGTCCGGTGGCAAGCTTTTCG	302
OY	308	atttccttgcccagaagaagafgcactgaatgtgagggagagccacaagcttcaatacgca	367
Db	303	ACTTCTCTGCGGAGAGAAGAGTCTATTGAGAGGAGATGAGGCCACGCAAGTTCCTCAACAAA	362
OY	368	tcctggatcgagggtgaactaccttcacacaaaagaattgtcacttlgtatctcaagccag	427
Db	363	TCCTTAGAGCGGTGTCACACTACCTGCACCTCCAAAGCGCATCGCACACTTTGACCTGAAGCCG	422
OY	428	aaaacattaagtltgttagacaagaataltcccatlccacacalcaagcttgtatgcttgg	487
Db	423	AGAACATCATATGTGTGCTGAGACAAGACAGCAGCAGCCCCCGCATTAAGCTCAATCGACTTGG	482
OY	488	gtctgcgcgcagaagaatagaagaatgagtagttaaataatatlttttgscgcggagat	547
Db	483	GCATCGCGCCACAGATCCAGCTGGACGCGAGTTCAAAACATCTTTGGCACCCCGAGT	542
OY	548	ttgtgtccccagaaatltgtaactacagacccttggtcgtgtgaggtctacatgtgtgagca	607
Db	543	TTGTGCGCCCGGAGATCGTAGAATATAGAGCCACTTGGCTTGAGAGCTGACATGTGTGAGCA	602
OY	608	taggcgtcatcaactlcaatcctctctaagtgtgagcaatcccttcccttggtgagaacgaagc	667
Db	603	TTGGGCTCATACCTCATACCTCCTCGACGGAGGCTCCCAATCTCTGGCGCAGAACCAAGC	662
OY	668	aggaaacctctggcaaatlcacatcagatgaggtiaagacttgaatgaagaatctcagcc	727
Db	663	AGGAGACCTCTACGACACATCTCAGCAGTAGAATACTGATTGATGAGGAAATCTTCAGCA	722
OY	728	atacgaagcagcttgcccaaggaacttatlctggaagcttctgtttaaagaagaccggaaac	787
Db	723	GCACGACGAGACTGGCCAAAGACTTCACCCGAGGCTGCTGTCAAAGACCCCAAGAGA	782
OY	788	ggct 791	
Db	783	GGAT 786	
RESULT	8		
LOCUS	B1046467/c	649 bp	mRNA linear EST 14-JUN-2001
DEFINITION	MNR-FN0209-070201-010-f01 FN0209 Homo sapiens cDNA, mRNA sequence.		
VERSION	B1046467		
KEYWORDS	B1046467.1 GI:14453089		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Dias Neto,E., Garcia Correa,R., Vertovskii-Almida,S., Britones,M.R., Ngai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed protein tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
MEDLINE	20202663		
COMMENT	Contact: Simpson A.J.G.		

|||||
Db 325 GAGCTTTGAGAACGACGACGCGTGCATCTCTGAGCTAGTGTGAGAGAG 384
Oy 301 ccttcgattccttgccaggaagagctagtgaggaaggagccagcttcatt 360
Db 385 CTCTTCTGACTTCTGCTGCGCAGAGGAGTCTGAGTGAAGAGGAGCGACCTTCATT 444
Oy 361 aagcagaatcctgagtgagggtgaactctctacacagaagaattctcattgctc 420
Db 445 AACGACATCTCGATGGGGGGAACCTTCATGCGCAAGAAATTCCTCACTTGATCTC 504
Oy 421 aagcagaagaacatlatgttcttagacaagaatattccattccacacatcaagctgatt 480
Db 505 AACCGAAGAAACATTAATGTTGTTGGACAGAAATATTCCTTCACACATCAAGCTGATT 564
Oy 481 g 481
Db 565 g 565

RESULT 13

LOCUS BF168866 999 bp mRNA linear EST 30-OCT-2000
DEFINITION 6017535F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4016911 5',
mRNA sequence.
ACCESSION BF168866
VERSION BF168866.1 GI:11049218
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 999)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9265 row: a column: 08
High quality sequence stop: 624.
Location/Qualifiers
1. 999

FEATURES
source

/organism="Mus musculus"
/strain="C57BL/6J (f1)"
/db_xref="taxon:10090"
/clone_lib="IMAGE:4016911"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/note="Organ: Lung; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 261 a 278 c 297 g 162 t
ORIGIN

Query Match 40.0%; Score 431.6; DB 10; Length 999;
Best Local Similarity 93.0%; Pred. No. 3.2e-89;
Matches 452; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Oy 1 atggagccattcaagcagagaagtgtagactttatgacatcgagagagctctggg 60
Db 109 ATGAGACGCTTCAACAGCAGAGGTGAGAGACTTTATGACATCGAGAGAGCTGGGC 168

Oy 61 agtggccagtttgccatctgtgaagaagtgcgaggaagaagcaggggttgatgca 120
Db 169 AGTGGCCAGTTTTCCTATCGTAAGAACTGCCGGAGAGACACAGGCTGGAGTATGCA 228
Oy 121 gccaaagttcaagaagcggcagagcggcgagccggcggtgtgagccggagagag 180
Db 229 GCCAAGTCAATTAAGAAAGAGCAGAGAGCGGGCCAGCCGTCGGGGCGGTGCCGGAGGAA 288
Oy 181 atcgaagcggagagtgagacatctcgccgaggtgtctgacacacaatgcatcagctgac 240
Db 289 ATCGAGCGGAGGTGAGCATCTCGCGCAGTCTGACCCCAACATCAATCAGCTGCAC 348
Oy 241 gacgtcatgagaccgcaccagcagctggtgacatcctttagttagtctgagagagag 300
Db 349 GACGCTCTATGAGAACCGACCGACGAGTGTCTATCTTACCTTAGCTAGTGTCCGAGGAA 408
Oy 301 ccttcgattccttgccaggaagagctagtgaggaaggagccacacattcatt 360
Db 409 CTGTTTGATTCTCGGCGCAGAGAGAGTCTTAAGTGAGAGAGAGACACACACTTCATT 468
Oy 361 aagcagaatcctgagtgagggtgaactctctacacagaagaattctcattgattc 420
Db 469 AACGACATCTCGATGGGGGGAATTAATTCCTTCACACAAAGAACTTGCTCACTTGATCTC 528
Oy 421 aagcagaagaacatlatgttcttagacaagaatattccattccacacatcaagctgatt 480
Db 529 AACCGAAGAAACATTAATGTTGTTGGACAGAAATATTCCTTCACACATCAAGCTGATT 588
Oy 481 gattt 486
Db 589 GGAATT 594

RESULT 14

LOCUS BE733365 1127 bp mRNA linear EST 15-SEP-2000
DEFINITION 601567911F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3842667 5',
mRNA sequence.
ACCESSION BE733365
VERSION BE733365.1 GI:10147357
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1127)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLM536 row: e column: 04
High quality sequence stop: 723.
Location/Qualifiers
1. 1127

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3842667"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCGAGGAG(5). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 268 a 305 c 357 g 197 t

ORIGIN

Query Match 39.8%; Score 429.8; DB 10; Length 1127;
Best Local Similarity 73.8%; Pred. No. 8.8e-83;
Matches 559; Conservative 0; Mismatches 197; Indels 1; Gaps 1;

```

Oy 1 atggagccattcaagcaagaagtgagagccttttgcacatcgagagagagctggg 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 ATGTCCAGCTTCAGCAGAGAGAGCTGAGACCTTTATGATGGGAGAGAGCTGGC 207
Oy 61 agtggccagtttgcacatcgtagaagaagtgcgggagagagacgggcttgagtagca 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 ACCGGCCAGTTTGCATCTGCGGAGAGTCCCGGAGAGGCGGAGAGAGATACGA 267
Oy 121 gccaaagttaacaagaagggagagccggggagccggcggtgtgagccggagagag 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 GCCAAGTTTATCAAGAGAGCGCGCTGTATCTACACCGCGCTGGGTGAGCCGGAGAG 327
Oy 181 atcgaagcggagagtgagcaacctgcggcaagtgtcgcacacacatgcatcagctgac 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 ATCGAGCGGAGAGTGAAATCTCGCGGAGATCCGGACCCCAACATCATCACCTGCAC 387
Oy 241 gacgtatagaagaacgcgcagcagtggtgcacatccttgcagctgagctgagagagag 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 GACATCTTGAAGAAAGAGAGAGAGGAGTGTCTCATCTGAGACTGTCTCTGCGGGAG 447
Oy 301 ccttcgattcctgcggccagagagagtgacactgagtgagagagagccacacatcatt 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 CTCTTGTACTTCTCGCGGAGAGAGAGTGTCTGACGAGAGAGAGAGAGAGAGAGAG 507
Oy 361 aagcagatcctggatgggtgtaactccttcacacacacacacacacacacacacacac 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 AAGCAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
Oy 421 aagcagacacacacacacacacacacacacacacacacacacacacacacacacac 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 568 AAGCCGGAAGAAATCATCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
Oy 481 gacttgcctgcgcacagaaatagaagatgagtgaaatgaagaatttttggagag 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 GACTTGGGAGATCGCGACAAAGATCGAGCGGGAGAGAGAGAGAGAGAGAGAGAGAG 687
Oy 541 cgggaatttgcctgcacacacacacacacacacacacacacacacacacacacacac 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 688 CCGGAGTTTGTGGCCCAAGAGATTGTGAATAGAGACCGCTGGGCTGGAGGCGGAGATG 747
Oy 601 tgaagatagagcgtcatcacctacatcctcttaagtggagacatcccttctctggagac 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 748 TGGAGCATCGGTGTCAATCACTATCTCTGAGGCTGAATTCCTGCTGACCAAGA 807
Oy 661 acgaagcagaagaacactcgggaatatacatcagtagtagtagtagtagtagtagtagtag 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 CCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866
Oy 721 ttcaagcatagcagcagcgtgcgcaggaagtagcttattc 757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 867 TTAGGAACCAACACTAAGTTGGCGAGAGAGTTGATCC 903

```

RESULT 15

BF727181

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BF727181 671 bp mRNA linear EST 05-JAN-2001
by17706.y1 Human Lens cDNA (Un-normalized, unamplified): By Homo
sapiens cDNA clone by17706 5', mRNA sequence.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 671)
Wistow, G.J., Bernstein, S., Behal, A. and Smith, D.
NEIBANK: EST analysis and bioinformatics for ocular genomics
Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 17 row: h column: 06
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers

FEATURES

source

1..671
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="by17706"
/clone_11b="Human Lens cDNA (Un-normalized, unamplified):
By"

/issue-type="Lens"
/dev-stage="Adult"
/lab-host="EMDH10B"
/note="Organ: Eye; Vector: pCMVSPORT2. Two human lenses
from different adults (both approximately 40 years old)
together yielded 200g of total RNA and 150ng mRNA for cDNA
library synthesis. A directionally cloned cDNA library in
the pCMVSPORT2 vector was constructed at Life Technologies
, essentially following the protocols of the Superscript
Plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
15'-pGATGTTTCTGATCGCGAGCGCGCC(17)15-3'. Not I/Bln
end inserts were cloned into the Not I/EcoR V sites in the
vector. Est analysis was performed on the unamplified
library at the NIH Intramural Sequencing Center (INSC)."

BASE COUNT

159 a 197 c 201 g 113 t 1 others

Query Match 39.6%; Score 428.2; DB 10; Length 671;
Best Local Similarity 77.6%; Pred. No. 1.7e-88;
Matches 517; Conservative 0; Mismatches 149; Indels 0; Gaps 0.

```

Oy 74 ccactgtgaagaagtcgggagagagcagggcttgagtagatgcagcgaattcatca 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 CGATGTCGGAGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 65
Oy 134 aagaagcagagacggcgagacggcggtgtgtagccggagagagatcagcggagag 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 AGAAGCGCGCTGTCTATCCAGCCGCGGTGTGAGCGGAGAGATCGAGCGGAGAG 125
Oy 194 tagacatcctgcggcaggtgtcgcacacacacacacacacacacacacacacacacac 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 TGAACATCTCTGGGAGAGATCGCGACCCCAACATATATACCTCGACGACATCTTGANA 185
Oy 254 accgcaacagcgtgtgtagacatccttgagctagtgctgtagagagagagctctcgatttc 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 ACNAGAGGAGAGTGTCTCATCTGAGACTGTGCTGTGCGGAGAGCTTTGACTTCC 245
Oy 314 tggccagaagaagtagtcaatcgtgtagagagagcagcagctcatcagcagatcctgg 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 TGGCGGAGAGAGAGTGTCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
Oy 374 atggagtgagacatcctcacaagaagaattgtcctcattgatcctcaagcagaagaaca 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 ACAGCGTTTACTTACTGCACTTAAGCGCATGCGACACTTGTAGCTGAAGCGGAGAGACA 365
Oy 434 ttatgttgtagaagaatattccatcaccacacacacacacacacacacacacacacacac 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 366 TCATGCTGCTGACAGACAGACGTCGCCAACCACGAAATCAAGCTCATCGACTTCGGCATCG 425
QY 494 CTCACGAAATAGAGAGATGAGATCgaatttaagaatatcttgggagcgcggaattgttg 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 CGCACAAGATCGAGCGGGAACGAGTTCAAGAACATCTTCGGCACCCCGAGTTGTGG 485
QY 554 CTCAGAAATLtgtaacttaacgagccctgggtctgagagctgaacatgtgagcataagcg 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 CCCAGAGATTGTGAAGTATGAGCCGCTGGAGGCGGACATGTGAGCATCGGTG 545
QY 614 tcaatcaactcaatccctcttaagtgagagcatccctctcttgaggagacagaaagagaaa 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 TCATCACTTATCTCTCTGAGCGGTGCATCNCGGTTCCTGGGGAGAACCAAGCAGGAGA 605
QY 674 caatggcaaatatcaatcatgagtgagtgtaagacttgatgaggaattcttcagccatacga 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 CGCTCACCAACATCTCAGCCGTGAAGTACGACTTTCGACGAGAGTACTTTAGCAACACCA 665
QY 734 gcgagc 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 666 GCGAGC 671
```

Search completed: May 17, 2002, 23:36:27
Job time: 5486 sec

This Page Blank (uspto)